Defense Technical Information Center
Compilation Part Notice

ADP013509

TITLE: Development of Performance and Effectiveness Metrics For Mechanical Diagnostic Technologies

DISTRIBUTION: Approved for public release, distribution unlimited

This paper is part of the following report:

TITLE: New Frontiers in Integrated Diagnostics and Prognostics. Proceedings of the 55th Meeting of the Society for Machinery Failure Prevention Technology. Virginia Beach, Virginia, April 2 - 5, 2001

To order the complete compilation report, use: ADA412395

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP013477 thru ADP013516

Development of Performance and Effectiveness Metrics For Mechanical Diagnostic Technologies

Rolf F. Orsagh Michael J. Roemer	Christopher J. Savage	Kathy McClintic
Impact Technologies 125 Tech Park Drive Rochester NY 14623	NAVSEA 5001 South Broad Street Philadelphia, PA 19112	Penn State ARL P.O. Box 30 State College, PA 16804

Abstract: In recent years, numerous anomaly detection and diagnostic technologies have been developed for various military and industrial applications to aid in the detection and classification of developing faults. In many cases, significant reductions in machinery total ownership costs have been achieved through the judicious application of these technologies. However, there is currently no consistent methodology available for assessing both the technical and economic benefits of these machinery diagnostic technologies. In response to this need, a virtual test bench is under development by the Navy for assessing the performance and effectiveness of machinery diagnostic systems. The test bench utilizes a 'plug 'n play' interface that can readily accept standardized diagnostic/prognostic tools and link them to real and model-based transitional data from appropriate condition based maintenance (CBM) platforms. The assessment process relies on a standard set of mathematical ground rules and a statistical framework to directly identify confidence bounds, robustness measures, and various diagnostic thresholds associated with specific mechanical diagnostic technologies. Specific performance and accuracy of the diagnostic algorithms at the component or subsystem level are evaluated with performance metrics, while system level capabilities in terms of achieving the overall operational goals of the diagnostic system will be evaluated with effectiveness measures. This qualification and validation methodology is utilized to compare a variety of diagnostic tools that are commonly used to analyze gearbox vibration.

Key Words: Diagnostics; Prognostics; Metrics; Diagnostic Qualification; Diagnostic Validation

Introduction: The US Navy's operational goals include improving mission readiness, and crew safety while reducing the support requirements and costs associated naval platforms. To accomplish these objectives the Navy is adopting condition based maintenance (CBM) practices. CBM is based on the principle of monitoring the condition of machinery and repairing it just prior to failure or an unacceptable level of performance degradation. Mission readiness can be enhanced by CBM through the elimination of unnecessary preventive maintenance and by identifying impending failures so that corrective action can taken in an efficient manner. CBM procedures can also protect crewmembers by identifying impending machinery malfunctions with sufficient warning to avert a catastrophic failure. By avoiding unnecessary preventive maintenance and allowing a scheduled response to impending failures, CBM can reduce the support requirements and total ownership cost associated with many types of machinery.

The success of a CBM program in a given application depends to a great extent upon the availability of useful diagnostic and prognostic information. CBM practices are most beneficial when maintenance actions can be planned well in advance, and corrective measures are carried out just prior to failure. Such precise maintenance scheduling can only occur through the use of timely and accurate diagnostic, or better yet, prognostic information. However, a consistent methodology for evaluating the technical and economic benefits of mechanical machinery diagnostic technologies does not currently exist. In response to this need, a virtual test bench is

under development by the Navy for assessing the performance and effectiveness of machinery diagnostic systems.

Performance Metrics Development

The performance of specific detection and diagnostic algorithms or subsystems of a CBM system are measured with Performance Metrics[i]. The functionality of these diagnostic algorithms or subsystems directly contributes to the overall effectiveness of the entire system. However, the ability to assess the accuracy and robustness of particular algorithms is often more straightforward when the technologies making up the system are checked separately. Also, from a design and development point of view, it is often more logical to work on the improvements to specific algorithms or processes at the elemental level rather than the overall systems level. Metrics of performance for diagnostic/prognostic algorithms or subsystems are arranged into three categories (detection, isolation, and prognosis) as shown in Figure 1. Detection metrics measure the ability of diagnostic tools to correctly classify machinery operation as either normal or anomalous. Isolation metrics measure the ability of diagnostic tools to accurately identify the root cause and corrective action for a fault. Prognosis metrics measure the ability of prognostic systems to accurately forecast the future condition of a mechanical system. Scores from the individual performance metrics are combined according the hierarchy to produce summary scores for each category, and for overall performance.

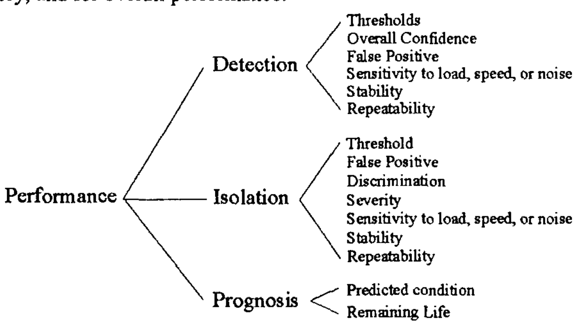


Figure 1 Performance Metrics

The ability of diagnostic/prognostic systems to detect and isolate faults or to predict failures is measured as a function of the fault severity. Figure 2 shows the confidence level reported by a hypothetical diagnostic tool and the corresponding fault severity level as functions of time. This could be the confidence that an anomaly exists or the confidence in a particular diagnosis. Varying operating conditions or noise could cause fluctuations in the diagnostic confidence level. The success function of the diagnostic tool is defined as the relationship between the average confidence and the average severity level. Note that this relationship may be used to assess either Boolean (0 or 1) confidence levels or continuous confidence levels within the same interval. The success function for the hypothetical diagnostic tool is plotted in Figure 3.

Fault severity must be established by objective and irrefutable measures to ensure that the assessments based upon it are accurate and impartial. This measure of severity will hereafter be referred to as the *ground truth severity level.* The ground truth severity of a system's condition may be assessed in a laboratory setting through the use of appropriate instruments or frequent inspections by nondestructive evaluation (NDE) techniques. Measurements of the fault severity are mapped onto the ground truth severity scale where zero represents a healthy operating condition, one represents an unacceptable level of performance degradation. Once the ground

truth is established, the anomaly detection threshold, isolation threshold, fault severity, stability, repeatability, and duty sensitivity metrics may be determined.

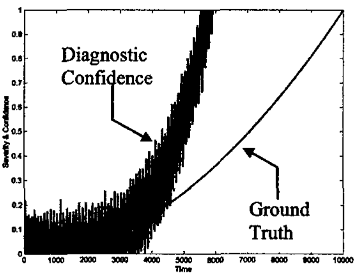


Figure 2 Diagnostic and Ground Truth Information

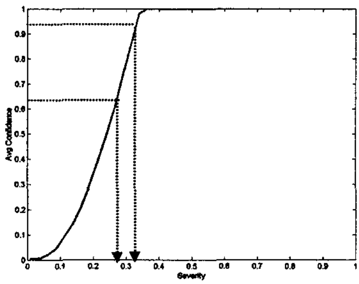


Figure 3 Success Function

Detection metrics

The ability of a diagnostic algorithm or overall system to detect anomalous machinery operating behavior is the most fundamental requirement for machinery health monitoring tool. For a diagnostic system to be useful it must detect anomalies associated with incipient faults so that corrective action may be taken in an efficient and timely manner. The *Detection Threshold Metric* measures a diagnostic algorithm or system's ability to identify anomalous operation associated with incipient faults with a specified confidence level. This metric is defined as the minimum ground truth severity corresponding to a designated confidence level on the detection success function as shown in Figure 3. Confidence levels of 67% and 95% corresponding to one and two standard deviations are used to calculate the detection threshold metric. Eq. (1) is used to calculate the detection threshold metric score.

$$\text{Detection Threshold} = 1 - S(c) \qquad (1)$$

where: S(c) = ground truth severity at a confidence of c

An assessment of the detection confidence level over the entire severity range for 0 to 1 is achieved with the *Overall Detection Confidence* metric defined in Eq (2). Graphically, the overall confidence score represents the area under the success function. An algorithm that detects an incipient fault with high confidence will receive a high Overall Confidence score, while an algorithm that does not report a fault until it becomes very severe would receive a low score.

$$OverallConfidence = \int_0^1 C(s)ds \qquad (2)$$

where: C(s) = The success function
s = severity

A confidence level that fluctuates wildly is difficult to interpret and therefore undesirable. For example, a diagnostic tool that produces a Boolean result of either no fault or fault may flicker as the fault severity approaches the detection level. The *Stability Metric* measures the range of confidence values that occur over the fault transition by integrating the peak to peak difference at each point on the success function as stated in Eq.(3).

$$\text{Stability} = 1 - \int_0^1 (C_H(s) - C_L(s))\ ds \quad (3)$$

where: $C_H(s)$ = maximum value of the success function at severity s
$C_L(s)$ = minimum value of the success function at severity s
s = severity

Ideally, diagnostic systems should detect anomalies over the full range of operating (duty) conditions such as loads, speeds, etc. The *Detection Duty Sensitivity Metric* measures the difference between the success functions of a diagnostic tool under two duty conditions as stated in Eq.(4).

$$DutySensitivity = 1 - \sqrt{\int_0^1 (C_1(s) - C_2(s))^2 ds} \quad (4)$$

where: $C_1(s)$ = success function at duty condition 1
$C_2(s)$ = success function at duty condition 2
s = severity

A diagnostic tool that incorrectly reports anomalies is unacceptable because it reduces availability and increases maintenance costs for the equipment. The *False Positive Confidence Metric* measures the frequency and upper confidence limit associated with false anomaly detection by a diagnostic tool. Calculation of the false confidence metric is based on the false positive function that is stated in Eq.(5) and an example is shown in Figure 4.

$$F(c) = n(c)/N \quad (5)$$

where: n(c) = number of false positive detection events with confidence > c
N = number of opportunities to detect a normal operating condition

Integration of the false positive function with respect to the confidence yields two parameters for assessing false anomaly detection by a diagnostic tool. The first parameter, α, represents the frequency of false positive anomaly detections and can be visualized as the area under the false positive function. The second parameter, β, is the confidence corresponding to 95% of α as shown in Figure 5. The mean value of the two parameters, α and β, helps determine the false confidence metric as shown in Eq.(6).

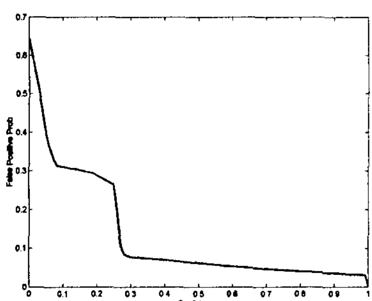


Figure 4 False positive function

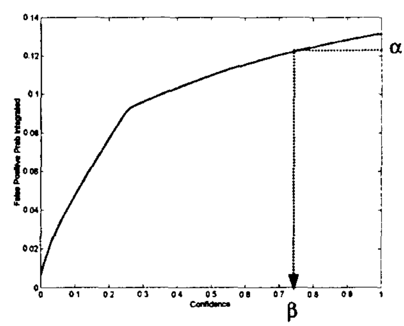


Figure 5 Integrated false positive function

$$FalseConfidence = 1 - \frac{\alpha + \beta}{2} \tag{6}$$

In an operational environment sensor data is sometimes contaminated with noise that may interfere with the operation of diagnostic algorithms. The robustness of an algorithm to noisy data is measured by the *Noise Sensitivity Metric.* Two aspects of the diagnostic system's response, change in the success function and increase in the false positive score are combined to form the noise sensitivity metric. The difference between the success functions of a diagnostic tool when the sensor data is contaminated with two different levels of noise. Eq. (7).

$$NoiseSensitivity = \left(1 - \sqrt{\int_0^1 (C_1(s) - C_2(s))^2 ds}\right) * (\Delta FalsePositive) \tag{7}$$

where: $C_1(s)$ = success function under noise condition 1
$C_2(s)$ = success function under noise condition 2
s = severity

Calibration of the performance metrics determine the weight that each individual metric carries in the category and overall composite scores. These weighting factors should reflect the specific requirements of the intended application, and therefore must be determined on a case by case basis. For example, when evaluating a gearbox diagnostic tool, knowledge of the gearbox's criticality (such as the main drive on helicopter vs. a redundant shipboard system) would determine the relative weight assigned to the detection threshold metric and the false confidence metric. The process of selecting weighting factors may be simplified by allowing the user to select a standard weighting scheme from a previously defined set or create a custom weighting scheme from scratch. A weighted average is used to calibrate and combine the individual performance metrics at the category level, and the category scores into an overall performance score as shown in Eq.(8).

$$CompositeScore = \frac{w_1 M_1 + w_2 M_2 + w_3 M_3 + \ldots + w_n M_n}{\sum w_i} \tag{8}$$

where: Mi = metric scores
w_i = weight assigned to metric i

Effectiveness Metrics
The overall effectiveness of a diagnostic system in terms of achieving the desired CBM goal is measured with Effectiveness Metrics[i]. This could include the integration of all the monitoring and diagnostic systems on the entire platform or a single diagnostic system made up of several different diagnostic algorithms. In either case, the effectiveness metrics utilize many of the same metrics as defined for the performance metrics. However, the resulting scores of the metric may be calibrated and combined differently based on the scope of their application. Some metrics such as cost, speed, complexity, robustness, and resource requirements are unique to the overall effectiveness of the diagnostic system and are therefore only defined as effectiveness metrics.

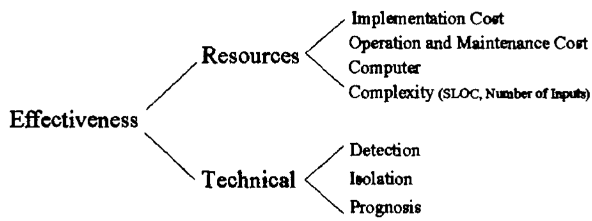


Figure 6 Effectiveness Metrics

Acquisition and implementation costs of the diagnostic system may have a significant effect on the system's cost effectiveness. The _Implementation Cost Metric_ simply measures the cost of acquiring and implementing a diagnostic system on a single application. If the diagnostic system is applied to several pieces of equipment, any shared costs are divided among them. Operation and maintenance costs may also play a significant role in determining whether a diagnostic system is cost effective. The _O&M Cost Metric_ measures the annual cost incurred to keep the diagnostic system running. These costs may include manual data collection, inspections, laboratory testing, data archival, relicensing fees and repairs.

The ability of the diagnostic algorithms or system to be run within specified time requirements and on traditional computer platforms with common operating systems is important when considering implementation on multiple machinery platforms. Therefore, a metric that takes into account computational effort as well as static and dynamic memory allocation requirements is necessary. The _Computer Resource Metric_ computes a score based on the normalized addition of CPU time to run (in terms of floating point operations), static and dynamic memory requirements for RAM and static source code space, and static and dynamic hard disk storage requirements. Computer requirements may be a significant issue in some applications such as aircraft.

Complex systems are generally more susceptible to unexpected behavior due to unforeseen events. The _System Complexity Metric_ measures the complexity of diagnostic systems in terms of the number of source lines of code (SLOCs) and the number of inputs required.

The individual effectiveness metric scores are combined to form an overall effectiveness score by means of a cost function. The benefits achieve through anomaly detection, fault isolation, and failure prediction are weighed against the costs associated with false alarms, inaccurate diagnoses, licensing, and resource requirements of implementing and operating a diagnostic tool. The simplified cost function in Eq. (9) states the _Technical Value_ provided by a diagnostic system for a given fault. The value of a diagnostic tool in a particular application is the summation of the benefits it provides over all the failure modes that it can diagnose less the implementation cost, operation and maintenance cost, and consequential cost of incorrect assessments as stated in Eq.(10).

$$Value = P_f * (D * \alpha + I * \beta) - (1 - P_f) * (P_D * \phi - P_I * \theta) \qquad (9)$$

where:

P_f = Probability (time-based) of occurrence for a failure mode
D = Overall Detection Confidence metric score
α = Savings realized by detecting a fault prior to failure
I = Overall isolation confidence metric score
β = Savings realized through automated isolation of a fault

P_D = False positive detection metric score
ϕ = Cost associated with a false positive detection
P_I = False positive isolation metric score
θ = Cost associated with a false positive isolation

$$TotalValue = \sum_{FailureModes} TechnicalValue_i - A - O - (1 - P_c) * \delta \quad (10)$$

where:
A = Acquisition and Implementation Cost
O = Life Cycle Operation and Maintenance Cost
P_c = Computer Resource Requirement score
δ = Cost of a standard computer system

CBM Metrics Database

One of the most significant aspects associated with the development and implementation of diagnostic system metrics is having well-documented fault data sets. Initial fault/failure data sets were obtained primarily from previously acquired test bed (including accelerated loading and run to failure tests) and simulation data sets with actual in-service data being applied later in the program. The Penn State ARL Mechanical Diagnostics Test Bed (MDTB) was utilized in this program as the basis for the diagnostic system metrics evaluation, testing and verification. The MDTB represents a wealth of well-documented data sets and information on gear, shaft and bearing faults and failures critical to Naval aircraft carrier day-to-day operations. The database of fault scenarios already developed under existing Multi-disciplinary University Research Initiative (MURI) provided an excellent basis and source of data from which the fault data sets utilized in this program were built upon. Identified metrics that require additional or more specific seeded fault or failure test data sets can be acquired from this test bed configuration or Penn State ARL's other test beds (Bearing Test Rig, Diesel Enhanced MDTB) throughout and after the duration of this program.

The metrics evaluation process is currently being implemented within the framework of a Test Bench that will utilize this database of sensor data from carefully constructed tests of selected CBM platforms as a basis for evaluating diagnostic/prognostic systems. Each test documents the transition of a mechanical system from a normal operating condition to failure or significantly degraded performance. Use of transitional data is necessary for the assessment of diagnostic/prognostic tools that rely on trending, and for evaluating the response of diagnostic/prognostic algorithms as a function of fault severity. Potential future sources for data of this type include the manufacturer of the equipment, Naval laboratories, and independent testing facilities. Contributions to the database should be screened to ensure data integrity and that the data remains unbiased toward any particular diagnostic/prognostic approach. The review process should include Naval engineers who will use the Test Bench to evaluate diagnostic/prognostic tools, and Naval maintenance officers who possess an intimate knowledge of the machinery reliability issues in the fleet.

Specifics of the MTBD Test Bed at ARL

The MDTB at Penn State was built as an experimental research station for the study of fault evolution in mechanical gearbox and power transmission components. It consists of a motor, gearbox, shafts, bearings, and a generator on a rigid steel platform. Gearboxes, shafts and bearings are instrumented with 52 sensors including accelerometers, thermocouples, acoustic emission sensors, and oil debris sensors. Tests are run at various load and speed profiles while

logging measurement signals for later analysis. Duty cycle profiles can be prescribed for any speed and load.

CBM Metrics Test Bench Web Application

Implementation of a standardized process and associated metrics for efficiently evaluating CBM information systems could potentially enhance the quality of diagnostic/prognostic technologies in two ways. First, doing so will allow the Navy and other users of diagnostic/prognostic tools to select the most appropriate algorithms for their application and verify the advertised capabilities of candidate systems. Second, developers of diagnostic systems may use the metric-based evaluation process to assess and improve their algorithms. To encourage participation, developers will have the option to evaluate their algorithms without creating any permanent record of the results.

In order to provide easy access to the CBM metrics developed under this program, a WEB-based prototype application called the **CBM Metrics Test Bench** has been developed to evaluate diagnostic technologies. Users of the site will upload algorithms to the server for evaluation and an e-mail will be issued to them indicating that their results are complete. The site will also provide access to a limited set of the maintained databases. However, a comprehensive set of data will only be accessible to Naval and other relevant DOD personnel for official use in qualification and validation of diagnostic tools.

On the Log-in page shown in Figure 7, the user can access the "Motivation and Evaluation Criterion", the "New User Registration", and the "User Log-In" links. Users who are not registered to use the web-site, may do so by clicking the "New User Registration" link. After successfully logging in, users may choose links that will allow them to obtain data, submit an algorithm, or view the results of an evaluation. Some of the transitional machinery failure data used in the evaluation will be available to facilitate the development of algorithms. Users will be able to download sample data sets from the web-site, or request a full data set to be mailed to them.

Figure 7 Log-in page

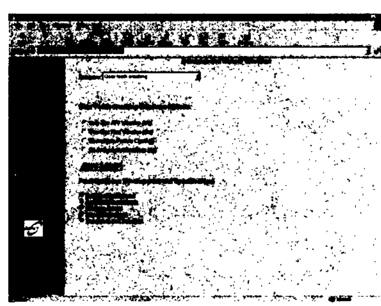


Figure 8 Database and sensor selection

To submit an algorithm, users will begin by uploading it to the server by either typing in the path and file name of the file containing their algorithm, or select it using "Browse". An algorithm description field is provided to allow users to identify their algorithms. After entering a file name, the algorithm will be assigned a Job ID that will be used to identify the algorithm within the Test Bench. Users may also choose the platform, faults, and sensor data on which their algorithm is evaluated. As the database grows, the user will be able to select a variety of failure modes for each platform. Information about the conditions under which each data set was collected is available through the links under the heading "Database Development and Specifications".

The weighting factors that are used to combine and calibrate the metric scores are accessible to the user on the metric weighting page shown in Figure 9. Users may view the definition of each metric by clicking on its name. When the user is satisfied with their choices, they may choose to perform the evaluation on either an official or a confidential basis. Algorithms that are evaluated on an official basis will have their scores added (anonymously) to a publicly accessible database.

Evaluation results are accessible on two levels. The lower level shows the scores earned by an algorithm while evaluating one particular fault on a platform. Users may view the definition of each metric by clicking on its name. The higher level results page presents the combined results for the algorithm against all of the selected faults. In the case of performance metrics, the scores are averaged, and for the effectiveness metrics reflect the sum of the technical values achieved by the algorithm for each fault type.

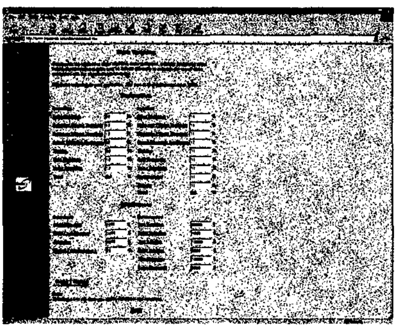


Figure 9 Metric weighting page

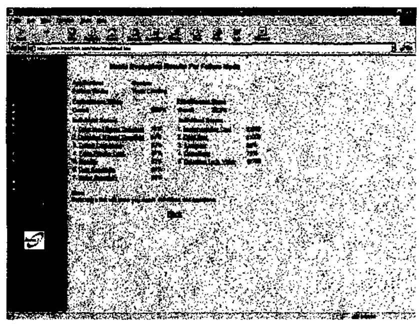


Figure 10 Evaluation results page

Results

The CBM Metrics Test Bench was used to evaluate the performance of ten anomaly detection algorithms for a gearbox. Gearbox failure data collected on the MDTB was used to evaluate the ability of the selected algorithms to detect gear tooth failures. During the test, cyclic loads as high as three times the rated load for the gearbox accelerated gear tooth failure rates. All of the algorithms utilize the same time domain vibration data, but process it in different ways.

Table 1 shows selected scores for each of the algorithms. For all of the metrics, a low score indicates an undesirable result, and high score indicates a desirable result. For example, a high Computer resource requirement score is awarded to algorithms that use a small portion of the computer's resources. Calculation of Detection Technical Value, Overall Performance, and Overall Effectiveness are based on weighting factors described in Eqs (8), (9), and(10). The factors used to calculate these results are stated in Tables 2 and 3. Evaluations of three diagnostic algorithms (RMS, Wavelet, and FM4) are described in detail.

Table 1 Metric Scores

Metric	RMS	Kurt	Wavelet	FM0	NA4	M6A	Dempster Shafer	FM4
Detection 1σ Threshold	22	19	27	68	64	77	73	76
Detection 2σ Threshold	0	0	12	52	19	64	64	64
Overall Confidence	51	39	44	75	64	84	79	82
False Positive Conf.	44	57	99	53	99	60	92	87
Stability	36	45	55	52	48	75	81	82
Duty Sensitivity	47	60	74	73	59	75	78	84
Noise Sensitivity	95	99	100	98	97	96	99	98
Implementation Cost $	1500	1500	1500	2000	2000	2000	2500	2000
O&M Cost $	700	700	700	1000	1000	1000	1000	1000
Computer	99	99	88	65	65	65	47	65
Complexity	99	99	97	79	79	79	78	79
Detection Tech. Value $	3255	2551	4387	6010	6374	7076	7640	7798
Overall Performance	42	46	61	65	66	75	82	82
Overall Effectiveness $	1052	333	1820	1021	1325	1879	1433	2448

Table 2 Performance Weighting Factors

Metric	Weight
Detection 1σ Threshold	10 %
Detection 2σ Threshold	10 %
Overall Confidence	20 %
False Positive Conf.	20 %
Stability	20 %
Duty Sensitivity	10 %
Noise Sensitivity	10 %

Table 3 Effectiveness Weighting Factors

Factor	Weight
Probability of Fault	20%
Cost of False Alarm	$4000
Benefit of Detection	$50000
Cost of Std. Computer	$2000

RMS is a simple and commonly used technique for detecting anomalous machinery operation. The RMS based algorithm calculates the root mean square value of the time domain vibration signal. The RMS level of a signal x consisting of N samples is calculated using Eq.(11). Figure 11 shows the diagnostic confidence reported by the RMS algorithm as compared to the ground truth severity level. The low performance scores assigned to RMS reflects the fact that RMS does not respond well in the early stages of gear damage and that the RMS level increases significantly with load. However, the low costs and low complexity (high complexity score) of the RMS algorithm make its overall effectiveness comparable to more sophisticated algorithms.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i} \quad (11)$$

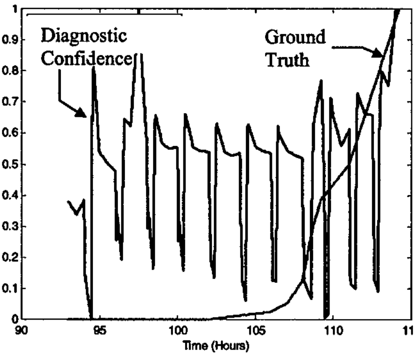


Figure 11 Diagnostic confidence reported by the RMS algorithm

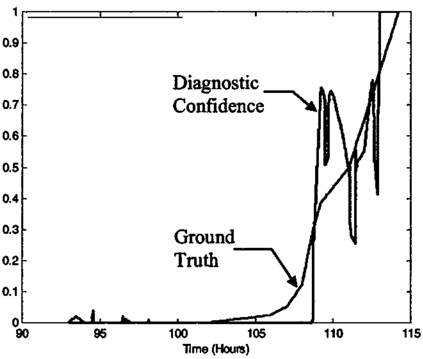


Figure 12 Diagnostic confidence reported by the Wavelet algorithm

The Wavelet algorithm uses a wavelet transform to analyze the nonstationary characteristics of vibration signal. The continuous wavelet transform of a time function *f(t)* is defined in Eq. (12) where *g(t)* is a given "mother wavelet" wavelet. The Morlet wavelet was chosen for *g(t)* and is defined mathematically by Eq. (13).

$$G(a,s) = |a|^{-1} \int_{-\infty}^{\infty} f(t) g[(t-s)/a] dt \qquad (12)$$

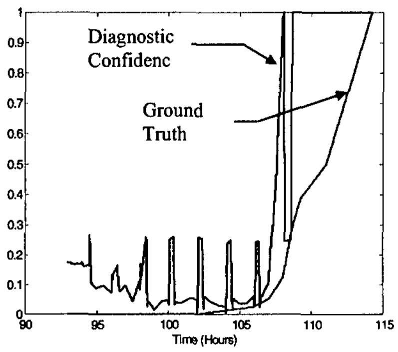
$$g(t) = \exp(-j\omega_0 t) \exp(-t^2/2) \qquad (13)$$

Where t is time and ω_0 is the fundamental (radian) frequency of the wavelet. Eq. (13) shows that the (complex) Morlet wavelet can be interpreted as a "modulated Gaussian." The actual Morlet wavelet chosen for the analysis is given by $\omega_0 = 5$ in Eq. (13) above. An adaptive IIR thresholding/tracking filter for processing wavelet output (at 550 Hz.) was also introduced. This kind of filter design is particularly robust against false alarms. The features resulting from the CWT processing include the number of detection counts (threshold crossings), and the peak amplitude and frequency obtained by a peak search of the CWT power spectral density near the frequency of interest (usually one of the shaft frequencies).

Figure 12 shows the diagnostic confidence reported by the Wavelet algorithm as compared to the ground truth severity level. Inspection of the Wavelet's diagnostic confidence will confirm that it warrants the high False Positive Detection score that it received. Furthermore, Wavelet shows very little load dependence as indicated by the Duty Sensitivity metric score.

The FM4 based algorithm uses the difference signal to detect changes in the vibration pattern resulting from damage on a limited number of teeth[ii]. FM4 is calculated for a difference signal d consisting of N samples according to Eq. (14). shows the diagnostic confidence reported by the FM4 algorithm as compared to the ground truth severity level. After calculating FM4, an empirical load correction was applied to reduce the load-induced fluctuations in the output. As a result of the load correction, the Duty Sensitivity metric score is higher (indicating that the confidence reported by the corrected algorithm is less dependent on the applied load. The same load correction technique was also applied to the M6A and Dempster Shafer (fusion) algorithms,

but not to the others. As expected, these load-corrected algorithms receive the highest duty sensitivity scores.



$$FM4 = \frac{N\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^4}{\left[\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2\right]^2} \quad (14)$$

Figure 13 Diagnostic confidence reported by the FM4 algorithm

Conclusion:

The metric-based process developed during this program clearly demonstrates the feasibility and potential benefits of a comprehensive system for evaluating the performance and effectiveness of diagnostic/prognostic tools. The principal achievements include the development and verification of diagnostic system metrics for evaluating and comparing the benefits advertised by system developers, and the eventual demonstration of these metrics in the assessment of various diagnostic tools. These achievements have been demonstrated through a comprehensive and easy-to-use internet-based software tool. The next necessary steps must include demonstration of the metrics software capabilities for various machinery diagnostic applications.

References:

[i] Essawu, M.A., Zein-sabatto, S, "Measures of Effectiveness and Measures of Performance for Machine Monitoring and Diagnostic Systems", Maintenance and Reliability Conference, May 1999.
[ii] Lebold, M., McClintic, K., Campbell, R., Byington, C., Maynard, K., "Review of Vibration Analysis Methods for Gearbox Diagnostics and Prognostics", 54th Meeting of the MFPT, May 2000

SMART SENSORS

Chair: Dr. Kang B. Lee
National Institute of Standards & Technology